

**APPLICATION FOR UNITED STATES LETTERS PATENT FOR
TRANSMISSION OF IMMERSIVE VIDEO VIA EXISTING VIDEO
INFRASTRUCTURE**

| <u>INVENTORS</u> | <u>Citizenship</u> | <u>Residence City and State</u> |
|--------------------|--------------------|---------------------------------|
| Furlan, John L. W. | USA | Belmont, CA |
| Fluker, Derek | USA | San Jose, CA |
| Hoffman, Robert G. | USA | Fremont, CA |



EL 595 180 516 US

TRANSMISSION OF IMMERSIVE VIDEO VIA EXISTING VIDEO INFRASTRUCTURE

Background of the Invention

Field of the Invention

5 [0001] This invention relates to the field of video signal transmission.

Background

[0002] Existing television infrastructure provides a means for capturing television video at a remote site (for example, using a remote unit), transferring the television video to an intermediate site (for example, a broadcast studio) for transmission to receiver sites (via broadcast, cable, 10 satellite, Internet, or similar technology).

[0003] Standard television video comprises interlaced lines of image information that combine to produce the visual effect of motion. Existing television infrastructure defines a frame composed of two fields of lines to effectuate the interlacing of lines in the frame. Television video presents a view of a scene that is captured by a video camera. The view of the scene captured by the camera is a function of the lens on the camera and the direction that the camera is pointed into the scene.

[0004] Immersive video comprises a stream of frames that allows a viewer to specify the view into the scene that is to be presented. An immersive video stream comprises a sequence of frames containing a wide-angle image of the scene (in some cases 360-degrees surrounding the lens, in other cases from a wide-angle lens such as 150-degree lens or a fish-eye lens). The immersive 20 video stream contains information beyond that provided by a normal view into the scene. Thus, a viewer can select which portion of the immersive video to view. There are a number of camera/lens technologies that capture immersive video frames. These include technologies that use a lens to capture an annular image of the scene around the lens, and those that use two or more wide-angle (often fisheye) lenses to capture hemispherical views of the scene around the lenses. 25 The multiple-lens technologies gather light that can be received by multiple cameras (or in some cases, by a single camera receiving images through both of the lenses). These technologies all capture warped images of the scene. Once the viewer specifies the viewpoint into the scene, the portion of the warped images that correspond to the view must be unwarped to present the undistorted view desired by the viewer.

[0005] Frames in immersive video streams do not have the same characteristics as standard television video. For example, an immersive video camera with a catadioptric lens that gathers light from 45 degrees above and below the horizon line has an aspect ratio of 4:1, that represents a 360-degree wide by 90-degree tall panorama (the aspect ratio will be 3.4:1 if the gathered light is from 45-degrees above and 60-degrees below the horizon line). Standard television video frames have an aspect ratio of 4:3 and consist of 640 by 480 pixels for the NTSC format and 704 by 512 for the PAL format. Another difference is that the amount of image data in a frame of an immersive video stream is much larger than the data in a standard frame of video. Immersive video frames generally are not interlaced (although they could be).

5 [0006] Immersive videos are currently sent across the Internet and generally are compressed for transmission using a compression/decompression mechanism (codec). The immersive video is stored on a server and made available to viewers on a network (such as a computer network, the Internet, or possible future broadcast networks).

[0007] One problem with providing live immersive video is that the immersive video is often captured at a site that is not local to a broadcast station or server farm. Thus, the live immersive video needs to be delivered to the broadcast station and/or server farm. The existing television infrastructure does not provide a cost effective way to deliver "live" immersive video from a remote site.

20 [0008] Traditionally, a remote television video stream is gathered by a remote unit at the camera/news/sport/event site, transmitted to a television studio where it is edited, possibly recorded, and then transmitted for viewing at receiver sites. The remote units are able to send standard television video to the television studio by using cable, microwave links, satellite, or other currently existing television infrastructure. In addition, the standard television video stream can be compressed (by a codec) for delivery over a network and the compressed video (or group of videos 25 compressed for different bandwidth utilization) sent to a server farm for delivery to clients.

30 [0009] One way to send an immersive video from the remote site is to have it compressed by a codec at the remote unit so that the data in the immersive frame fits within a television video frame. This requires a codec at each remote unit and one at the studio. Codecs that can process the frame rate and resolution required by immersive video are expensive (either in hardware cost or in the computer capability required to execute a software codec at video rates). Currently, remote units

generally do not include such codecs. Thus, adding such a codec to the remote unit (for example, in each roving television station van) increases the cost of the remote unit. In addition, because the images captured through the camera/lens are generally not rectangular (usually circular, multiple circular, or annular) standard compression algorithms used by the codecs are not as efficient as if
5 the image were rectangular.

[0010] United States Patent No.: 5,280,540, *Video Teleconferencing Systems Employing Aspect Ratio Transformation*, dated 1/18/1994 by Addeo et al. teaches means for transmitting a 16:9 aspect ratio image using a 4:3 aspect ratio transmission frame. However, Addeo does not teach or suggest the problems addressed by the current invention nor the approach taken by the inventors to
10 solve these problems.

[0011] Because immersive video does not have the same characteristics as standard television video, existing television infrastructure is not well suited for transmitting immersive video from a remote unit to the broadcast station.

[0012] It would be advantageous to be able to format the immersive video stream within a standard video stream so that existing and future television infrastructure/technology can be used to send an immersive video stream to a designated site where the immersive video can be converted into a deliverable form (for example, by broadcast or Internet service).

Summary of the Invention

[0013] The problems associated with sending an immersive video using existing television infrastructure are addressed by aspects of the inventions disclosed herein. In one preferred embodiment, an immersive video is acquired at a first location, packed into one or more standard television frames and sent to a second location using standard television infrastructure.
20

[0014] Another preferred embodiment receives at least one standard television video frame that contains an immersive video frame, unwarps a portion of the immersive video frame into a view
25 and presents the view.

[0015] Yet another preferred embodiment includes the steps of acquiring an immersive video frame, packing the immersive video frame into at least one standard television video frame that is sent to a second location using television infrastructure to be received at a television receiver where

a portion of the immersive video frame within the standard television video frame is unwarped into a view and presented.

[0016] Still other preferred embodiments include apparatus for sending and/or receiving such immersive videos using television infrastructure, and of systems for doing the same.

5 [0017] In addition, another preferred embodiment is of computer program products that cause a computer to perform the operations of these and similar apparatus and systems.

[0018] The foregoing and many other aspects of the present invention will no doubt become obvious to those of ordinary skill in the art after having read the following detailed description of the preferred embodiments that are illustrated in the various drawing figures.

10 Description of the Drawings

[0019] Fig. 1 illustrates a field of view of a catadioptric lens in accordance with a preferred embodiment;

[0020] Fig. 2A illustrates an annular image that represents the field of view of Fig. 1;

[0021] Fig. 2B illustrates a panoramic representation of the annular image of Fig. 2A having an aspect ratio of 4:1;

[0022] Fig. 2C illustrates a panoramic representation of the annular image of Fig. 2A having an aspect ratio of 3.4:1;

[0023] Fig. 2D illustrates a real world three-dimensional environment including a warped circular image resulting from a wide-angle lens in accordance with a preferred embodiment;

[0024] Fig. 2E illustrates a frame containing dual hemispherical images of a scene in accordance with a preferred embodiment;

[0025] Fig. 2F illustrates a frame containing a projection resulting from the dual hemispherical images of Fig. 2E in accordance with a preferred embodiment;

25 [0026] Fig. 3 illustrates an immersive video transmission architecture in accordance with a preferred embodiment;

[0027] Fig. 4A illustrates a first packing of a 4:1 warped representation in accordance with a preferred embodiment;

[0028] Fig. 4B illustrates a second packing of a 4:1 warped representation in accordance with a preferred embodiment;

5 [0029] Fig. 4C illustrates a first packing of a 3.4:1 warped representation in accordance with a preferred embodiment;

[0030] Fig. 4D illustrates a second packing of a 3.4:1 warped representation in accordance with a preferred embodiment;

10 [0031] Fig. 4E illustrates a split-frame packing of a warped representation in accordance with a preferred embodiment;

[0032] Fig. 4F illustrates a packing of dual scaled hemispherical images in accordance with a preferred embodiment;

[0033] Fig. 4G illustrates a split-frame packing of a dual hemispherical image in accordance with a preferred embodiment;

[0034] Fig. 4H illustrates truncated packing of an hemispherical image in accordance with a preferred embodiment;

[0035] Fig. 5 illustrates an immersive video transmission process in accordance with a preferred embodiment;

20 [0036] Fig. 6 illustrates an immersive video receiver process in accordance with a preferred embodiment;

[0037] Fig. 7 illustrates a second immersive video transmission process in accordance with a preferred embodiment;

[0038] Fig. 8 illustrates a second immersive video receiver process in accordance with a preferred embodiment; and

25 [0039] Fig. 9 illustrates a viewing process in accordance with a preferred embodiment.

Description of the Preferred Embodiments

Notations and Nomenclature

[0040] The following ‘notations and nomenclature’ are provided to assist in the understanding of the present invention and the preferred embodiments thereof.

- 5 [0041] **Procedure** — A procedure is a self-consistent sequence of computerized steps that lead to a desired result. These steps are defined by one or more computer instructions. These steps can be performed by a computer executing the instructions that define the steps. Thus, the term “procedure” can refer (for example, but without limitation) to a sequence of instructions, a sequence of instructions organized within a programmed-procedure or programmed-function, or a
10 sequence of instructions organized within programmed-processes executing in one or more computers. Such a procedure can also be implemented directly in circuitry that performs the required steps.

Description

[0042] Fig. 1 illustrates a field of view **100** captured by a catadioptric lens attached to a camera. All light intersecting a viewpoint **101** is captured on either side of an horizon line **103** for a substantially 360-degree band-of-light within a vertical field of view **105** defined by a first angle **107** above the horizon line **103** and a second angle **109** below the horizon line **103**. The first angle **107** and the second angle **109** need not (but can) have the same value (for example both angles being 45-degrees or the first angle **107** being 45-degrees and the second angle **109** being 60-degrees.
15
20

[0043] Fig. 2A illustrates an annular image **200** that represents the field of view **100** of a catadioptric lens. The annular image **200** can be unwrapped by designating an edge **201** and mapping the annular image **200** into a panorama.

- 25 [0044] Fig. 2B illustrates a panoramic image **210**, that results from unwrapping the annular image **200** of Fig. 2A, and that has an aspect ratio of 4:1. The 4:1 aspect ratio results from the first angle **107** and the second angle **109** each having a value of 45-degrees.

[0045] Fig. 2C illustrates a panoramic image **220** that has an aspect ratio of 3.4:1 that results from the first angle **107** having a value of 45-degrees and the second angle **109** having a value of 60-degrees.

[0046] The aspect ratio of the panoramic band of light captured by a catadioptric lens is determined by comparing the vertical field of view 105 with 360-degrees. Thus, if the first angle 107 and the second angle 109 were both 45-degrees the aspect ratio would be 4:1. However, if the first angle 107 was 45-degrees and the second angle 109 was 60-degrees the aspect ratio would be
5 3.4:1.

[0047] **Fig. 2D** illustrates a real world three-dimensional environment 250 that has been imaged by a wide-angle lens 251. The real world three-dimensional environment 250 can be defined by the cartesian coordinate system in X, Y and Z with the viewpoint defined to be the origin of the coordinate system. One skilled in the art will understand that the real world three-dimensional environment 250 can also be defined using spherical or cylindrical or other coordinate systems.
10 The viewing direction of the user, as determined from the user's input, can be given as a viewing vector in the appropriate coordinate system. An image plane 253 containing a warped wide-angle image 255 can be defined by a two dimensional coordinate system in U and V, with the origin of the coordinate system coincident with the origin of the X-Y-Z coordinate system. If the field of view of the wide-angle lens 251 is sufficient, and the lens is rotationally symmetric about the viewing axis, the warped wide-angle image 255 will be substantially circular in the U-V plane.

[0048] **Fig. 2E** illustrates a frame of hemispherical images 260 that contains a first hemispherical image 261 and a second hemispherical image 263. Each of the hemispherical images result from capturing a substantially back-to-back 180-degree field of view through a very-wide-angle lens (for example a fish-eye lens) such that both hemispheres, when combined, provide a 360-degree by 180-degree image. A camera system capable of capturing the frame of hemispherical images 260 is described in United States Patent 6,002,430, *Method and Apparatus for Simultaneous Capture of a Spherical Image*. Another arrangement for capturing a 360-degree image is described in United States patent 5,796,426, *Wide-Angle Image Dewarping Method and Apparatus*.
25

[0049] **Fig. 2F** illustrates a rectangular representation 270 that shows one result of mapping two hemispherical images such as shown in Fig. 2E onto a full panoramic image having an aspect ratio of 2:1. For a full panorama, the opposite edges of the rectangular representation connect.

[0050] Some of these immersive video frames provide enough information to create a complete 360-degree by 180-degree panorama or a 360-degree by 90-degree panorama. Others provide enough information for a partial panorama (for example, the frame shown in Fig. 4H).

[0051] **Fig. 3** illustrates an immersive video transmission architecture **300** used to transmit immersive videos of a scene taken by a video camera **301** equipped with a warping lens **303**. Each of the immersive video frames from the video camera **301** contains a warped representation of the scene around the warping lens **303**. Each frame acquired from the video camera **301** is communicated to a remote broadcast unit **305** by either wire or a wireless communication channel. The remote broadcast unit **305** can be equipped with a communications link **307** (for example, but without limitation a satellite link, a microwave link, or any other television signal transmission mechanism). One skilled in the art will understand that the immersive video can be gathered (for example, but without limitation) by a digital video camera, an analog video camera in communication with a digitizer, a video playback device, or a computer.

[0052] The warping lens **303** can be one or more wide-angle lenses (including fish-eye lenses and rectilinear lenses) and/or one or more catadioptric lens.

[0053] The warped representation of the scene that results when the warping lens **303** is a catadioptric lens is a complete or partial annular image. Other types of lenses produce warped representations having characteristics particular to the lens type. For example, a fish-eye lens produces a circular representation of a hemispherical portion of the scene. A wide-angle lens is another lens that will produce a warped representation. In addition, a rectilinear wide-angle lens can be used to capture a perspective-corrected image of the scene that is less warped.

[0054] In the remote broadcast unit **305**, each frame of the immersive video is processed by a video processing device **309** to map the warped image (such as by unwrapping the annular image **200**) into at least one standard television video frame. This standard television video frame is then sent using the television signal transmission mechanism. Fig. 3 for example, illustrates a satellite communication system where the signal is first sent to a satellite **311** that re-transmits the signal to a broadcast facility **313** where it is received by a television signal receiver mechanism **315** that is connected to a computer **317**. Once the standard television video frame is received, the computer **317** can use a codec **319** to compress the immersive video frames into compressed frames for

200
199
198
197
196
195
194
193
192
191
190
189
188
187
186
185
184
183
182
181
180
179
178
177
176
175
174
173
172
171
170
169
168
167
166
165
164
163
162
161
160
159
158
157
156
155
154
153
152
151
150
149
148
147
146
145
144
143
142
141
140
139
138
137
136
135
134
133
132
131
130
129
128
127
126
125
124
123
122
121
120
119
118
117
116
115
114
113
112
111
110
109
108
107
106
105
104
103
102
101
100
99
98
97
96
95
94
93
92
91
90
89
88
87
86
85
84
83
82
81
80
79
78
77
76
75
74
73
72
71
70
69
68
67
66
65
64
63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
47
46
45
44
43
42
41
40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1

storage on a server computer 321. The server computer 321 can then make the immersive video available for streaming.

[0055] One skilled in the art will understand that the codec **319** is optional and that raw uncompressed video is often provided. Where the codec **319** is used, it can compress frames independently (such as a JPEG compression) and/or compress the frames for video streaming (such as an MPEG compression). Finally, such a one will understand that the codec **319** can be embodied as a specialized hardware device and/or as a computer executing codec software.

[0056] The server computer 321 can be connected to a computer network 323 (such as the Internet) and serves information from the stored frames (for example, compressed frames) to a client device 325. The client device 325 unwarps a portion of each frame it receives to present a viewer-designated view (for example, in real-time through a computer monitor or television set, by recording the view on a video tape, a disk or optical film, or on paper). In some embodiments, the client device 325 sends viewpoint information to the server computer 321 so that the server computer 321 will generate the view and send the view to the client device 325 for presentation (or to provide bandwidth management such as described by US Patent Application 09/131,186). The server computer 321 can also provide the compressed frames over a broadcast, cable, satellite, or other network for receipt by the client device 325.

[0057] In another preferred embodiment, once the broadcast facility 313 receives the video from the remote broadcast unit 305, the video can be compressed one or more ways by a streaming video encoder, (for example, RealProducer or Windows Media Encoder). The video can be compressed by different amounts to target a particular bandwidth required for streaming the video. The compressed video streams can then be provided to users by the broadcast facility 313 or provided to a server farm to make the video streams available.

[0058] In yet another preferred embodiment, a director or cameraperson at the broadcast facility 313 can use a computer to select a view from the immersive video and broadcast that selected view to television receivers (either as the primary picture or as a picture-in-picture view).

[0059] The client device 325 can be a client computer, a television receiver, a video conferencing receiver, a personal organizer, an entertainment system, a set-top-box, or other device capable of generating a view from the compressed frames received over a network such as the computer network 323 or other transmission mechanism such as a microwave link, a television

cable system, a direct subscriber line (DSL) system, a satellite communication system, a fiber communication system, an Internet, a digital television system, an analog television system, a wire system, or a wireless system.

[0060] In another preferred embodiment, the standard television video frame is a high definition television (HDTV) video frame and the television infrastructure is capable of supporting HDTV transmission and reception.

[0061] Thus, an immersive video frame captured in real-time can be captured at a remote site, packed into a standard television video frame and transmitted to the broadcast facility 313 using existing television transmission infrastructure. The central site can then reconstruct the immersive video, compress the immersive video, make it available over a network, and/or select a view into the immersive video and broadcast the selected view. In addition, the compressed immersive video can be broadcast to a set-top-box for processing by the set-top-box to allow a viewer to select his or her own view.

[0062] Fig. 4A through Fig. 4E illustrate some of the ways an immersive video frame can be packed into at least one standard television video frame. **Fig. 4A** illustrates one way a warped representation (for example, a panoramic image having an aspect ratio of 4:1 captured by a catadioptric lens) can be apportioned to fit within a standard television video frame 400. In this example, when each half of the panoramic image is transformed into the standard television video frame 400, the transformation process also scales the vertical dimension of each half of the panoramic image so that both halves of the panoramic image (a first 180-degree scaled portion of the panoramic image 401 and a second 180-degree scaled portion of the panoramic image 403) can be stored in the standard television video frame 400 (leaving an unused portion of the standard television video frame 405). This approach maintains the resolution in the horizontal direction of the panorama at the expense of the resolution in the vertical direction.

[0063] In the case of an annular image, because the information of the annular image is less towards the center of the annular image than at the outer edge (and equivalently in the panoramic image version of the annular image), this approach can result in a loss of information in the vertical direction. Images from wide-angle lenses can also have distortions that affect the amount of information available to parts of an image and can have corresponding affects when the image is packed within the standard television video frame 400.

00
01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

[0064] One skilled in the art will understand that while the warped representation can first be mapped (for example, by unwrapping an annular image) into a panorama and the panoramic image then scaled to fit into the standard television video frame 400, the transformation from the warped representation to the standard television video frame 400 can also include the required scaling.

5 [0065] **Fig. 4B** illustrates another way that a 4:1 aspect ratio panoramic image frame can be apportioned into a standard television video frame 410. In this example, each half of the panoramic image is packed into the standard television video frame 410 by scaling the dimension when performing the transformation while maintaining the vertical dimension. In this example, when each half of the panoramic image is transformed into the standard television video frame 410
10 the transformation scales the horizontal dimension of the half panoramic representation so that both halves of the panoramic image (for example, a first scaled 180-degree portion of the annular image 411 and a second scaled 180-degree portion of the annular image 413) will fit in the standard television video frame 410 (leaving an unused portion of the standard television video frame 415). This approach maintains the information in the vertical direction of the panorama at the expense of the information in the horizontal direction.

0
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
0
20
30
40
50
60
70
80
90
0
10
20
30
40
50
60
70
80
90

[0066] **Fig. 4C** illustrates how a 3.4:1 aspect ratio frame can be packed into a standard television video frame 420. In this example, when each half of the panoramic image is transformed into the standard television video frame 420, the transformation scales the vertical dimension of the half panoramic image so that both halves of the panoramic image (for example, a first scaled 180-degree portion of the annular image 421 and a second scaled 180-degree portion of the annular image 423) will fit in the standard television video frame 420 (leaving an unused portion of the standard television video frame 425). This approach maintains the information in the horizontal direction of the panorama at the expense of the information in the vertical direction. Thus, for annular images this approach further reduces the available resolution in the vertical dimension.

25 [0067] **Fig. 4D** illustrates another way that a 3.4:1 aspect ratio frame can be packed into a standard television video frame 430. In this example, each half of the panoramic image is packed into the standard television video frame 430 by scaling the horizontal dimension when performing the transformation while maintaining the vertical dimension. In this example, when each half of the panoramic image is transformed into the standard television video frame 430 the transformation
30 scales the horizontal dimension of the half panoramic image so that both halves of the panoramic image (a first scaled 180-degree portion of the panoramic image 431 and a second scaled 180-

degree portion of the panoramic image 433) will fit in the standard television video frame 430. This approach maintains the information in the vertical direction of the panorama at the expense of the information in the horizontal direction. Thus, this approach is often preferred when used with annular images because it tends to maintain the resolution in the vertical direction. In addition, 5 substantially all of the standard television video frame 430 is packed with panoramic information.

[0068] The examples provided by Fig. 4A through Fig. 4D allow for transmitting immersive video frames using standard television broadcast infrastructure at the standard television frame rate (typically 30 frames-per-second). Often a frame rate of 15 fps is satisfactory for presentation of an immersive video. In this circumstance, the warped image can be transformed into two standard 10 television video frames. **Fig. 4E** illustrates a pair of standard television video frames 440 (a first standard television video frame 441 and a second standard television video frame 443) for transmitting the warped representation. The first standard television video frame 441 contains a first 180-degree portion of the panoramic image 445 and the second standard television video frame 443 contains the second 180-degree portion of the panoramic image 447. Each standard television video frame contains an unused portion of the standard television video frame 449. In addition, each standard television video frame contains a portion that tags whether that frame is a first partial frame or a second partial frame. Thus, a first indicator portion 451 identifies the first standard television video frame 441 to be the first partial frame while a second indicator portion 453 identifies the second standard television video frame 443 to be the second partial frame. In addition, each partial frame can include a designated portion that contains other information (for example, a first ancillary data portion 455 and a second ancillary data portion 457). The first ancillary data portion 455 can be used to pass additional information from the remote site to the broadcast facility. This information can be sent as text for display, or as binary information encoded into the frame. One skilled in the art will understand that the first indicator portion 451 and the second indicator portion 453 are generally positioned in substantially the same area of their 25 respective standard television video frame (although this condition is not required). The first indicator portion 451 and the second indicator portion 453 are used to indicate frame ordering.

[0069] One skilled in the art will understand that techniques similar to the above can be applied to sequencing more than two frames.

30 [0070] Other header or tag information can be included in the ancillary data portion of the frame. This can include the size and orientation of partial frames; how many television frames are

used to assemble a panoramic frame; lens characteristics; error detection and correction codes; a frame rate value (so we can transmit sources in non-real-time or sources whose rate is not divisible by 30 fps (for example for PAL use (25fps)).

[0071] Future high-resolution cameras will allow higher resolution immersive video frames

5 (resulting in a panoramic image of 1920 X 480 pixels). These higher resolution frames can be packed into three standard video frames of 670 X 480 to provide a frame rate of 10 frames per second.

[0072] Where the warped image is obtained from one or more wide-angle lenses, the data making up the captured circular images can be equivalent to a panorama having a 2:1 ratio.

10 Scaling can be applied to the 2:1 panoramic view to pack the information into at least one standard television video frame as previously discussed. In addition, the hemispherical information can be stored in a standard television video frame as is without prior mapping to a panorama.

[0073] **Fig. 4F** illustrates a television frame containing scaled hemispherical images **460**. As previously discussed with respect to Fig. 2E and Fig. 2F two hemispherical images of a scene can be used to capture 360-degree by 180-degree information suitable for use in an immersive video. The television frame containing scaled hemispherical images **460** contains a first scaled hemispherical image **461** and a second scaled hemispherical image **463**. Each of these images is scaled in the horizontal dimension (with respect to the frame) so that the two images can fit within the standard television video frame **460**. The scaling of each image can be accomplished to retain the maximum amount of information (for example, by rotating the hemispherical image to maximize the retained information along the horizon line of the image).

[0074] **Fig. 4G** illustrates a pair of standard television video frames **470** that contain unscaled or uniformly scaled hemispherical images. The pair of standard television video frames **470** includes a first standard television video frame **471** and a second standard television video frame **473** that contain a first hemispherical image **475** and a second hemispherical image **477** respectively along with an unused portion **479**. Similar to the frames described with respect to Fig. 4E, the pair of standard television video frames **470** includes a first indicator portion **481**, a second indicator portion **483**, a first ancillary data portion **485**, and a second ancillary data portion **487** having similar functions as described with respect to Fig. 4E.

0071
0072
0073
0074
0075
0076
0077
0078
0079
0080
0081
0082
0083
0084
0085
0086
0087
0088
0089
0090
0091
0092
0093
0094
0095
0096
0097
0098
0099
0100
0101
0102
0103
0104
0105
0106
0107
0108
0109
0110
0111
0112
0113
0114
0115
0116
0117
0118
0119
0120
0121
0122
0123
0124
0125
0126
0127
0128
0129
0130
0131
0132
0133
0134
0135
0136
0137
0138
0139
0140
0141
0142
0143
0144
0145
0146
0147
0148
0149
0150
0151
0152
0153
0154
0155
0156
0157
0158
0159
0160
0161
0162
0163
0164
0165
0166
0167
0168
0169
0170
0171
0172
0173
0174
0175
0176
0177
0178
0179
0180
0181
0182
0183
0184
0185
0186
0187
0188
0189
0190
0191
0192
0193
0194
0195
0196
0197
0198
0199
0200
0201
0202
0203
0204
0205
0206
0207
0208
0209
0210
0211
0212
0213
0214
0215
0216
0217
0218
0219
0220
0221
0222
0223
0224
0225
0226
0227
0228
0229
0230
0231
0232
0233
0234
0235
0236
0237
0238
0239
0240
0241
0242
0243
0244
0245
0246
0247
0248
0249
0250
0251
0252
0253
0254
0255
0256
0257
0258
0259
0260
0261
0262
0263
0264
0265
0266
0267
0268
0269
0270
0271
0272
0273
0274
0275
0276
0277
0278
0279
0280
0281
0282
0283
0284
0285
0286
0287
0288
0289
0290
0291
0292
0293
0294
0295
0296
0297
0298
0299
0300
0301
0302
0303
0304
0305
0306
0307
0308
0309
0310
0311
0312
0313
0314
0315
0316
0317
0318
0319
0320
0321
0322
0323
0324
0325
0326
0327
0328
0329
0330
0331
0332
0333
0334
0335
0336
0337
0338
0339
0340
0341
0342
0343
0344
0345
0346
0347
0348
0349
0350
0351
0352
0353
0354
0355
0356
0357
0358
0359
0360
0361
0362
0363
0364
0365
0366
0367
0368
0369
0370
0371
0372
0373
0374
0375
0376
0377
0378
0379
0380
0381
0382
0383
0384
0385
0386
0387
0388
0389
0390
0391
0392
0393
0394
0395
0396
0397
0398
0399
0400
0401
0402
0403
0404
0405
0406
0407
0408
0409
0410
0411
0412
0413
0414
0415
0416
0417
0418
0419
0420
0421
0422
0423
0424
0425
0426
0427
0428
0429
0430
0431
0432
0433
0434
0435
0436
0437
0438
0439
0440
0441
0442
0443
0444
0445
0446
0447
0448
0449
0450
0451
0452
0453
0454
0455
0456
0457
0458
0459
0460
0461
0462
0463
0464
0465
0466
0467
0468
0469
0470
0471
0472
0473
0474
0475
0476
0477
0478
0479
0480
0481
0482
0483
0484
0485
0486
0487
0488
0489
0490
0491
0492
0493
0494
0495
0496
0497
0498
0499
0500
0501
0502
0503
0504
0505
0506
0507
0508
0509
0510
0511
0512
0513
0514
0515
0516
0517
0518
0519
0520
0521
0522
0523
0524
0525
0526
0527
0528
0529
0530
0531
0532
0533
0534
0535
0536
0537
0538
0539
0540
0541
0542
0543
0544
0545
0546
0547
0548
0549
0550
0551
0552
0553
0554
0555
0556
0557
0558
0559
0560
0561
0562
0563
0564
0565
0566
0567
0568
0569
0570
0571
0572
0573
0574
0575
0576
0577
0578
0579
0580
0581
0582
0583
0584
0585
0586
0587
0588
0589
0590
0591
0592
0593
0594
0595
0596
0597
0598
0599
0600
0601
0602
0603
0604
0605
0606
0607
0608
0609
0610
0611
0612
0613
0614
0615
0616
0617
0618
0619
0620
0621
0622
0623
0624
0625
0626
0627
0628
0629
0630
0631
0632
0633
0634
0635
0636
0637
0638
0639
0640
0641
0642
0643
0644
0645
0646
0647
0648
0649
0650
0651
0652
0653
0654
0655
0656
0657
0658
0659
0660
0661
0662
0663
0664
0665
0666
0667
0668
0669
0660
0661
0662
0663
0664
0665
0666
0667
0668
0669
0670
0671
0672
0673
0674
0675
0676
0677
0678
0679
0680
0681
0682
0683
0684
0685
0686
0687
0688
0689
0690
0691
0692
0693
0694
0695
0696
0697
0698
0699
0700
0701
0702
0703
0704
0705
0706
0707
0708
0709
07010
07011
07012
07013
07014
07015
07016
07017
07018
07019
07020
07021
07022
07023
07024
07025
07026
07027
07028
07029
07030
07031
07032
07033
07034
07035
07036
07037
07038
07039
07040
07041
07042
07043
07044
07045
07046
07047
07048
07049
07050
07051
07052
07053
07054
07055
07056
07057
07058
07059
07060
07061
07062
07063
07064
07065
07066
07067
07068
07069
07070
07071
07072
07073
07074
07075
07076
07077
07078
07079
07080
07081
07082
07083
07084
07085
07086
07087
07088
07089
07090
07091
07092
07093
07094
07095
07096
07097
07098
07099
070100
070101
070102
070103
070104
070105
070106
070107
070108
070109
070110
070111
070112
070113
070114
070115
070116
070117
070118
070119
070120
070121
070122
070123
070124
070125
070126
070127
070128
070129
070130
070131
070132
070133
070134
070135
070136
070137
070138
070139
070140
070141
070142
070143
070144
070145
070146
070147
070148
070149
070150
070151
070152
070153
070154
070155
070156
070157
070158
070159
070160
070161
070162
070163
070164
070165
070166
070167
070168
070169
070170
070171
070172
070173
070174
070175
070176
070177
070178
070179
070180
070181
070182
070183
070184
070185
070186
070187
070188
070189
070190
070191
070192
070193
070194
070195
070196
070197
070198
070199
070200
070201
070202
070203
070204
070205
070206
070207
070208
070209
070210
070211
070212
070213
070214
070215
070216
070217
070218
070219
070220
070221
070222
070223
070224
070225
070226
070227
070228
070229
070230
070231
070232
070233
070234
070235
070236
070237
070238
070239
070240
070241
070242
070243
070244
070245
070246
070247
070248
070249
070250
070251
070252
070253
070254
070255
070256
070257
070258
070259
070260
070261
070262
070263
070264
070265
070266
070267
070268
070269
070270
070271
070272
070273
070274
070275
070276
070277
070278
070279
070280
070281
070282
070283
070284
070285
070286
070287
070288
070289
070290
070291
070292
070293
070294
070295
070296
070297
070298
070299
070300
070301
070302
070303
070304
070305
070306
070307
070308
070309
070310
070311
070312
070313
070314
070315
070316
070317
070318
070319
070320
070321
070322
070323
070324
070325
070326
070327
070328
070329
070330
070331
070332
070333
070334
070335
070336
070337
070338
070339
070340
070341
070342
070343
070344
070345
070346
070347
070348
070349
070350
070351
070352
070353
070354
070355
070356
070357
070358
070359
070360
070361
070362
070363
070364
070365
070366
070367
070368
070369
070370
070371
070372
070373
070374
070375
070376
070377
070378
070379
070380
070381
070382
070383
070384
070385
070386
070387
070388
070389
070390
070391
070392
070393
070394
070395
070396
070397
070398
070399
070400
070401
070402
070403
070404
070405
070406
070407
070408
070409
070410
070411
070412
070413
070414
070415
070416
070417
070418
070419
070420
070421
070422
070423
070424
070425
070426
070427
070428
070429
070430
070431
070432
070433
070434
070435
070436
070437
070438
070439
070440
070441
070442
070443
070444
070445
070446
070447
070448
070449
070450
070451
070452
070453
070454
070455
070456
070457
070458
070459
070460
070461
070462
070463
070464
070465
070466
070467
070468
070469
070470
070471
070472
070473
070474
070475
070476
070477
070478
070479
070480
070481
070482
070483
070484
070485
070486
070487
070488
070489
070490
070491
070492
070493
070494
070495
070496
070497
070498
070499
070500
070501
070502
070503
070504
070505
070506
070507
070508
070509
070510
070511
070512
070513
070514
070515
070516
070517
070518
070519
070520
070521
070522
070523
070524
070525
070526
070527
070528
070529
070530
070531
070532
070533
070534
070535
070536
070537
070538
070539
070540
070541
070542
070543
070544
070545
070546
070547
070548
070549
070550
070551
070552
070553
070554
070555
070556
070557
070558
070559
070560
070561
070562
070563
070564
070565
070566
070567
070568
070569
070570
070571
070572
070573
070574
070575
070576
070577
070578
070579
070580
070581
070582
070583
070584
070585
070586
070587
070588
070589
070590
070591
070592
070593
070594
070595
070596
070597
070598
070599
070600
070601
070602
070603
070604
070605
070606
070607
070608
070609
070610
070611
070612
070613
070614
070615
070616
070617
070618
070619
070620
070621
070622
070623
070624
070625
070626
070627
070628
070629
070630
070631
070632
070633
070634
070635
070636
070637
070638
070639
070640
070641
070642
070643
070644
070645
070646
070647
070648
070649
070650
070651
070652
070653
070654
070655
070656
070657
070658
070659
070660
070661
070662
070663
070664
070665
070666
070667
070668
070669
070670
070671
070672
070673
070674
070675
070676
070677
070678
070679
070680
070681
070682
070683
070684
070685
070686
070687
070688
070689
070690
070691
070692
070693
070694
070695
070696
070697
070698
070699
070700
070701
070702
070703
070704
070705
070706
070707
070708
070709
070710
070711
070712
070713
070714
070715
070716
070717
070718
070719
070720
070721
070722
070723
070724
070725
070726
070727
070728
070729
070730
070731
070732
070733
070734
070735
070736
070737
070738
070739
070740
070741
070742
070743
070744
070745
070746
070747
070748
070749
070750
070751
070752
070753
070754
070755
070756
070757
070758
070759
070760
070761
070762
070763
070764
070765
070766
070767
070768
070769
070770
070771
070772
070773
070774
070775
070776
070777
070778
070779
070780
070781
070782
070783
070784
070785
070786
070787
070788
070789
070790
070791
070792
070793
070794
070795
070796
070797
070798
070799
070800
070801
070802
070803
070804
070805
070806
070807
070808
070809
070810
070811
070812
070813
070814
070815
070816
070817
070818
070819
070820
070821
070822
070823
070824
070825
070826
070827
070828
070829
070830
070831
070832
070833
070834
070835
070836
070837
070838
070839
070840

[0075] **Fig. 4H** illustrates a television frame **490** containing a truncated hemispherical image **491** that maximizes the information in the width direction and reduces an unused space **493** in the television frame **490** by sacrificing information in the vertical direction of the television frame **490**. The truncated hemispherical image **491** can result from a wide-angle lens (including a fisheye lens). The truncated hemispherical image **491** can also be treated as a warped, limited-angle panorama (as compared to the previously discussed panoramas that can extend for substantially 360-degrees). Thus, the television frame **490** contains a higher resolution, limited-angle panorama that still allows generation of a user-specified view into the panorama as is subsequently described.

[0076] **Fig. 5** illustrates an immersive video transmission process **500** that initiates at a ‘start’ terminal **501** and continues to an ‘initialization’ procedure **503** that performs any initialization in preparation for transmitting an immersive video from the remote broadcast unit **305** to the broadcast facility **313**. After initialization, the immersive video transmission process **500** continues to a ‘determine lens parameters’ procedure **505** that determines the characteristics of the warping lens **303** or lenses. Some of these characteristics can include the field-of-view, number of lenses and exposure information. This information can be determined from the images received by the video camera **301**, by prompting an operator for input, or by use of other mechanisms.

[0077] Once the lens parameters are determined, a ‘receive immersive video frame’ procedure **507** receives an immersive video frame from a stream of immersive video frames from the video camera **301** or playback unit (for example, a recorder/player or a storage device). A ‘transform immersive video frame’ procedure **509** apportions the received immersive video frame into at least one standard television video frame. This standard television video frame is transmitted to the broadcast facility **313** by a ‘transmit video frame’ procedure **511**. The immersive video transmission process **500** continues to the ‘receive immersive video frame’ procedure **507** to process the next video frame received by the video camera **301**. This process continues until there are no more immersive video frames to be received or until terminated by some condition (for example, termination by an operator).

[0078] **Fig. 6** illustrates an immersive video receiver process **600** that initiates at a ‘start’ terminal **601** and continues to an ‘initialization’ procedure **602** that performs any initialization in preparation for receiving the at least one standard television video frame of an immersive video sent by the ‘transmit video frame’ procedure **511** of Fig. 5. After initialization, a ‘receive video

frame' procedure **603** receives a standard television video frame that contains information representing the immersive video frame captured by the video camera **301**.

[0079] In some embodiments, a 'reconstruct immersive video frame' procedure **605** extracts each portion of the immersive video frame and regenerates the original panorama in memory. The 5 regenerated panorama can be the same scale as the original panorama, but need not be.

[0080] Other embodiments, that can serve the immersive video from the standard television video frame, need not perform the 'reconstruct immersive video frame' procedure **605** because the server and/or client software are enabled to process the immersive video directly from the 10 information in the standard television video frame without need for an intermediate regenerated panorama.

[0081] A 'save frame' procedure **607** stores the information received by the 'receive video frame' procedure **603** in computer memory, hard disk, or (when storing the received video frame) on videotape or other video storage mechanism. Once the frame is stored, the immersive video receiver process **600** continues to a 'video complete' decision procedure **609** that determines whether the video stream has ended or whether the immersive video receiver process **600** has been terminated. If the video stream has not ended, the immersive video receiver process **600** continues back to the 'receive video frame' procedure **603** to process the next frame. However, if the 'video complete' decision procedure **609** determines that the video stream has completed or that the process is to end, the immersive video receiver process **600** continues to a 'compress and store video' procedure **611**. The 'compress and store video' procedure **611** can compress the received video and stores either or both the uncompressed and compressed streams. This compression is accomplished by a codec device or codec software executing within a computer.

[0082] Once the video is stored, the immersive video receiver process **600** terminates through an 'end' terminal **613**.

[0083] One skilled in the art will understand that the immersive video receiver process **600** as previously described accumulates all the video frames before compressing them. Such a one will understand that other preferred embodiments allow (for example) every frame to be individually compressed, allow key frames to be compressed with subsequent non-key frames including difference information, or use streaming compression. These compression mechanisms can operate 30 (for example, but without limitation) after all the frames have been received, in parallel as each

frame is received, or in parallel on a set of received frames. Furthermore, although compression will generally be used, it is not required to practice the invention.

[0084] Fig. 7 illustrates an immersive video transmission process **700** that initiates at a ‘start’ terminal **701** and continues to an ‘initialize’ procedure **703** that performs any initialization in preparation for transmitting an immersive video from the remote broadcast unit **305** to the broadcast facility **313**. After initialization, the immersive video transmission process **700** continues to a ‘determine lens parameters’ procedure **705** that determines the characteristics of the warping lens **303** or lenses. Some of these characteristics can include the field-of-view, number of lenses and exposure information. This information can be determined from the images received by the video camera **301**, by prompting an operator for input, or by use of other mechanisms.

[0085] Once the lens parameters are determined, a ‘receive immersive video frame’ procedure **707** receives a warped representation from a stream of immersive video frames from the video camera **301**. A ‘transform 1/2 immersive video into first video frame’ procedure **709** transforms substantially half of the warped representation (if the video frame contains an annular image, half of the annular image is unwrapped) into a standard television video frame and marks the standard television video frame as a first partial frame by filling the first indicator portion **451** with a first identified signal such as a “white” color. The standard television video frame is then transmitted by the ‘transmit first video frame’ procedure **711**. If the half panorama fits with the standard television video frame, it need not be scaled.

[0086] The second portion of the warped representation is transformed by the ‘transform 1/2 immersive video into second video frame’ procedure **713** into a standard television video frame and marks the standard television video frame as a second partial frame by the second indicator portion **453** with a “black” color. The standard television video frame is then transmitted by a ‘transmit second video frame’ procedure **715**.

[0087] After the second partial frame is transmitted by the ‘transmit second video frame’ procedure **715** the immersive video transmission process **700** continues to the ‘receive immersive video frame’ procedure **707** to receive and process the next warped representation. The process continues until no additional immersive video frames are received by the ‘receive immersive video frame’ procedure **707** or until an event occurs (such as termination by an operator).

0
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95

[0088] One skilled in the art will understand that the first partial frame and the second partial frame are distinguished by differences between values in the first indicator portion **451** and the second indicator portion **453**. Such a one will also understand that the “white” and “black” colors only need to be distinguishable such that the receiver can determine which of standard television
5 video frame is the first partial frame and which is the second partial frame. In addition, such a one will understand that additional information can be included in the first ancillary data portion **455** and/or the second ancillary data portion **457** as the standard television video frame is being constructed. Finally, such a one will understand that the previously described techniques can be applied to more than two standard television video frames so long as the resulting immersive video
10 frame rate is satisfactory.

[0089] **Fig. 8** illustrates an immersive video receiver process **800** that initiates at a ‘start’ terminal **801** and initializes at an ‘initialize’ procedure **803**. Once the immersive video receiver process **800** has initialized, it continues to a ‘wait for first frame’ procedure **805** that receives frames sent using the immersive video transmission process **700** of Fig. 7 until it detects a first partial frame by examining the indicator portion of the standard television video frame for the first indicator portion **451**. Next, the immersive video receiver process **800** continues to a ‘receive first video frame’ procedure **807** that receives the standard television video frame that contains the first partial frame. A ‘receive second video frame’ procedure **809** then receives the standard television video frame that contains the second partial frame. Once both partial frames are received, a ‘reconstruct immersive video frame’ procedure **811** assembles the partial frames into a panoramic video frame that is saved by the ‘save immersive video frame’ procedure **813**. Furthermore, the ‘reconstruct immersive video frame’ procedure **811** can extract information stored in the first ancillary data portion **455** and/or the second ancillary data portion **457**.

[0090] One skilled in the art will understand that the assembly process can start on information received in the first video frame once the first video frame is received.
25

[0091] A ‘video complete’ decision procedure **815** determines whether the immersive video has completed. If not, the immersive video receiver process **800** continues to the ‘receive first video frame’ procedure **807** (some embodiments — those that have the possibility of losing synchronization — can return to the ‘wait for first frame’ procedure **805**) to receive and process the
30 next first partial frame and second partial frame.

[0092] Once the panoramic frames are saved, a ‘compress and store video on server’ procedure 817 optionally compresses the video frames and stores the compressed or non-compressed frames on a server. The immersive video receiver process 800 completes through an ‘end’ terminal 819

[0093] One skilled in the art will understand that the immersive video receiver process 800 as previously described accumulates all the video frames before compressing them. The invention was described in such a way as to make it more understandable. Such a one will understand that other embodiments allow every frame to be individually compressed, allow key frames to be compressed with subsequent non-key frames including difference information. These compression mechanisms can operate (for example, but without limitation) after all the frames have been received, in parallel as each frame is received, or in parallel on a set of received frames. In addition, one skilled in the art will understand that the functions of the immersive video receiver process 600 and the immersive video receiver process 800 can be combined to automatically detect whether partial frames or complete frames are being received.

[0094] **Fig. 9** illustrates a viewing process 900 that initiates at a ‘start’ terminal 901 and initializes at an ‘initialize’ procedure 903. The viewing process 900 continues to a ‘receive frame from server’ procedure 905 (for example, but without limitation, by using techniques such as those described in United States Patent 6,043,837). Once the frame is received, data from within the frame is unwarped to generate a view according to a user-specified viewpoint (for example, but without limitation, by using techniques such as those described in United States Patent 5,796,426). The view can be displayed for example on a computer monitor, a television, by being printed on a tangible media or otherwise presented to a viewer. In addition, the view can be recorded on optically sensitive film, a disk (such as a magnetic disk, CD or DVD), a videotape, or other tangible recording media.

[0095] One skilled in the art will understand that the one embodiment of the invention allows immersive video frames to be sent from a remote site to a receiving site using standard television infrastructure. Such a one will also understand that some of the many uses of the invention include live broadcast of sporting events, newscasts, and any other situation where a real-time immersive video is to be transferred from the remote broadcast unit 305 to the broadcast facility 313. Once at the broadcast facility 313 the immersive video can be compressed and provided for distribution to others by transmission from the broadcast facility 313 for viewer control on a set-top-box, by storage on a server for access over a computer network for viewer control on a computer, by

selecting a view from the immersive video at the broadcast facility **313** for separate broadcast or picture-in-picture inclusion within an existing broadcast.

[0096] From the foregoing, it will be appreciated that the invention has (without limitation) the following advantages:

- 5 [0097] 1) Removes the need for expensive codec devices at the camera site.
- [0098] 2) Uses existing television transmission infrastructure to send an immersive video from the camera site to a central site.
- [0099] 3) Removes the need for a high-data-rate communication link at the camera site.
- 10 [0100] 4) Removes the need for high-speed network connections between the remote broadcast unit **305** and the broadcast facility **313**.
- [0101] 5) Removes the need for streaming video expertise at the remote site as the streaming is done at the studio.
- 15 [0102] Although the present invention has been described in terms of the presently preferred embodiments, one skilled in the art will understand that various modifications and alterations may be made without departing from the scope of the invention. Accordingly, the scope of the invention is not to be limited to the particular invention embodiments discussed herein.

0096 0097 0098 0099 0100 0101 0102